Bayesian Sampling Plans with Interval Censoring

Wanbo Lu\textsuperscript{1,3}, and Tzong-Ru Tsai\textsuperscript{2}

\textsuperscript{1} Southwestern University of Finance and Economics, Chengdu, CHINA
\textsuperscript{2} Tamkang University, Tamsui, Taipei, TAIWAN
\textsuperscript{3} Corresponding author: Wanbo Lu, e-mail: luwb@swufe.edu.cn

Abstracts

This paper employs Bayesian approach to establish acceptance sampling plans for life tests with interval censoring. Assume that interval data have a multinomial distribution, and the interval probabilities are random and vary from lot to lot according to a conjugate prior of Dirichlet distribution. A Bayes risk is defined with a suitable loss function and a predictive distribution. Optimal Bayesian sampling plans are determined by minimizing the Bayes risk per lot. An example is used and some optimal Bayesian sampling plans with three equally-spaced intervals are tabulated for illustration. Sensitivity analysis are conducted to evaluate the influence of the parameter of prior distribution, the cost per sampled item and the cost per used unit time on the proposed Bayesian sampling plans.

Keywords: reliability, Dirichlet distribution, life test plan, loss function

1. Introduction

Acceptance sampling is one aspect of quality assurance in the applications of quality control. When quality characteristic is the lifetime of item, acceptance sampling plans are developed for lifetime data. Due to saving test time and cost, engineers would like to infer life information of components based on incomplete data through censoring tests, for instance type I censoring test, type II censoring test or hybrid censoring test. Among these censoring tests, type I censoring test is popular in practice because the termination time is known in advance. However, practitioners might only count the failure numbers at some fixed times and did not measure the exact lifetimes during a censoring test for the purpose of administrative convenience. We call such a censoring test as interval censoring test and the collected failure numbers are the interval censoring data or the grouped data.

In practice, an interval censoring test is easier to operate than conducting a type I censoring test, but the interval censoring data contain less information than that contained in type I censoring data. Some acceptance sampling plans have been developed with interval censoring data, see, for example, Ehrenfeld (1962), Kendall and Anderson (1971), Seo and Yum (1993), Chen and Mi (1998), Lu and Tsai (2009a), Lu and Tsai (2009b), and Tsai and Lin (2010).

In acceptance sampling studies, many acceptance sampling schemes have been investigated in the literatures. Among these, the decision theory approach is particularly attractive, more precise and scientific to develop sampling plans based on the economic consideration. The Bayesian methods arise naturally when prior information is available for planning and estimation. Fertig and Mann (1974), Padgett and Tsokos (1979), Lam (1988), Fernandez (2000), Lin et al. (2002), Zhang and Meeker (2005), Chen et al. (2007) investigated all kinds of Bayesian sampling plans (BSPs) with various loss functions.

Considering a suitable cost model, the paper employs Bayesian approach to establish acceptance sampling plans for grouped data. The rest of this paper is organized as follows. The development of BSPs for grouped data is discussed in Section 2. An example is used and some BSPs are tabulated for illustration in Section 3. In Section 4, some concluded remarks are given.
2. The Sampling Plans

Assume that items are manufactured from lot by lot each of size $N$. Before items are released for consumer use, they are submitted for inspection of acceptance sampling under an interval censoring test as the following: A sample of $n$ items are drawn randomly from a lot and suffer a life test simultaneously. Let $\tau_i=0$ and $\tau_i$ be the initial time and termination time of the test, respectively, and let $\tau_{k+1}=\infty$. The order of $k$ prescribed inspection times is specified as $0<\tau_1<\cdots<\tau_k<\tau_{k+1}$. At the time $\tau_i$, the number of failed items in the interval $[\tau_{i-1}, \tau_i]$ is observed as $X_i$. The decision rule $\delta$ is the probability of accepting the lot when $X_i$ is observed.

Let $\chi$ be the sample space of the lifetime at an item within the interval $(-\infty, \theta)$, if the action of accepting the lot is made, and $0$ otherwise. The action of rejecting the lot is $\delta(X_i)=\chi_i$, given $X_i$. Let $X=(X_1,\cdots,X_k)$ and $Y=(Y_1,\cdots,Y_k)$. The joint probability density functions of $X$ and $Y$ are presented, respectively as

$$P(x|\theta)=\frac{n!}{x_1!x_2!\cdots x_k!}\theta_1^{x_1}\theta_2^{x_2}\cdots\theta_k^{x_k}(1-\theta_1-\cdots-\theta_k)^{n-x_1-x_2-\cdots-x_k},$$

and

$$P(y|\theta)=\frac{(N-n)!}{y_{10}!y_{20}\cdots y_{k0}!}\theta_1^{y_{10}}\theta_2^{y_{20}}\cdots\theta_k^{y_{k0}}(1-\theta_1\cdots-\theta_k)^{N-n-x_{10}-x_{20}-\cdots-x_{k0}},$$

where $x_{k+1}=n-\sum_{i=1}^k x_i$, $y_{k+1}=N-n-\sum_{i=1}^k y_i$, $0<\theta_i\cdots<\theta_k\leq 1$, and $\sum_{i=1}^k \theta_i=1$; $\theta=(\theta_1,\theta_2,\cdots,\theta_k)$, where $\theta_i$ is the probability of the lifetime at an item within the interval $[\tau_{i-1}, \tau_i]$, $i=1,2,\cdots,k$. Because the manufacturing conditions may not be consistent, the mean life of items varies from lot to lot. It follows that the probabilities $\theta$ are random and follow a prior density $\pi(\theta)$.

Let $c_i$ be the cost per sampled item, and let $c$ be the cost per unit time used. Thus, $c_i n + c_r$ is the total cost for the life test. Let $\delta(X)$ be the action in favor of accepting the lot or not based on the sampled items $X$; namely, $\delta(X)=1$, if the action of accepting the lot is made, and $\delta(X)=0$, if the action of rejecting the lot is made. Let $\mathcal{X}$ be the sample space of interval censored sample $X$. The decision rule $\delta$ is a mapping from $\mathcal{X}$ into $[0,1]$ for each $x$ in $\mathcal{X}$. The $\delta(X)$ can be defined to be the probability of accepting the lot when $X=x$ is observed.

Let $y_0=(y_{10}, y_{20}, \cdots, y_{k0})$ be $k$ specified failure numbers for evaluating whether an un-sampled item is good or not. The determination of $y_0$ is given at the end of this section. Let $L_d(Y) = (N-n)I(\cup_{j=1}^k A_j)$, where $I(\cdot)$ is an indicator function. Let $A_i = \{Y_i \geq y_{i0}\}$, $A = \{Y < y_{10}, Y_1 < y_{20}, \cdots, \sum_{j=1}^i Y_j < y_{i0}, 0, \sum_{j=1}^k Y_j \geq y_{k0}\}$, $j=1,\cdots,k$. $L_d(Y)$ is the loss for all un-sampled items with random failure numbers $Y$ incurred in an accepted lot. Let $Y$ denote the loss per un-sampled item incurred in a rejected lot with $0<\gamma<1$. For each $(X,Y)$, the overall cost or loss associated with a sampling plan $(n,y_0,\delta)$ is defined as the following equation

$$L(n,y_0,\delta,X,Y) = c+ c_r n + c_r + (N-n)\gamma +(N-n)\delta(X) \{I(\cup_{j=1}^k A_j) - \gamma\}$$

The Bayes risk of a sampling plan $(n,y_0,\delta)$ is defined by

$$R(\pi, n,y_0,\delta) = E_x [E_Y L(n,y_0,\delta,X,Y) | \theta]$$

$$= c+ c_r n + c_r + (N-n)\gamma +(N-n)E_x [E_Y \delta(X) \{I(\cup_{j=1}^k A_j) - \gamma\} | \theta],$$

(1)

The probability density function (pdf) of $X$ is $P(x|\theta)=\int_0^\infty P(x|\theta)\pi(\theta)\theta d\theta$, and the posterior pdf of $\theta$, given $X=x$ is $\pi(\theta|x)=P(x|\theta)\pi(\theta)/P(x)$. Thus $P(x|\theta)\pi(\theta)=P(x)\pi(\theta|x)$. Also, let the predictive pdf of $Y$, given $X=x$ be $P(y|x)=\int_\theta P(y|\theta)\pi(\theta|x)\theta d\theta$, and let $H(x)=\sum_y I(\cup_{j=1}^k A_j)P(y|x)$. We can show that the Bayes risk of (1) can be rewritten as
\[ R(\pi, (n_x y_0, \delta)) = E_x E_{X,Y} L(n_x y_0, \delta, X, Y) \]
\[ = c n + c \tau + (N-n) y + (N-n) \sum_{y_j=0}^{n \alpha_j} \cdots \sum_{y_k=0}^{n \alpha_k} \delta(x) \{(H(x) - \gamma) P(x) \} \]
\[ (2) \]

The Bayesian decision rule can be determined by

\[ \delta_x(x | n y_0) = \begin{cases} 1, & \text{if } H(x) - \gamma < 0, \\ 0, & \text{otherwise.} \end{cases} \]
\[ (3) \]

Moreover, the optimal BSP, \( (n_x y_0, \delta_x) \), can be determined based on equation

\[ R(\pi, (n_x y_0, \delta)) = \min_{\lambda \in \Lambda} R(\pi, (n_x y_0, \delta(x | n y_0))) \]

We model \( \theta \) by the Dirichlet prior distribution with parameters \( \alpha, \ldots, \alpha_{k+1} \), its pdf is

\[ \pi(\theta) = \Gamma(\alpha) / \prod_{i=1}^{k+1} \Gamma(\alpha_i) \cdot \theta_1^{\alpha_1-1} \cdots (1-\theta_i-\theta_{i+1})^{\alpha_{i+1}-1} \]

where \( 0 \leq \theta_1, \ldots, \theta_k \leq 1 \) and \( \alpha = \sum_{j=1}^{k+1} \alpha_j \). Therefore,

\[ H(x) = \frac{\Gamma(\alpha) \Gamma(N-n+1)}{\prod_{i=1}^{k+1} \Gamma(\alpha_i) \Gamma(N+n+1)} \left\{ \sum_{y_j=0}^{n \alpha_j} \cdots \sum_{y_k=0}^{n \alpha_k} \prod_{i=1}^{k+1} \Gamma(y_i + \alpha_i) \left[ \prod_{i=1}^{k+1} \Gamma(y_i + 1) \right] + \sum_{y_j=0}^{n \alpha_j} \cdots \sum_{y_k=0}^{n \alpha_k} \prod_{i=1}^{k+1} \Gamma(y_i + 1) \right\} \]
\[ (4) \]

The average loss of accepting the lot without sampling can be found as

\[ K_a = N \left\{ \sum_{i=1}^{k} \left[ \int I(A_j | \theta) \pi(\theta) d\theta \right] \right\} \]
\[ (5) \]

The average loss of rejecting the lot without sampling can be found as \( K_r = N \gamma \).

Let \( K_0 = \min \{ K_a, K_r \} \); the lot is accepted without test if \( R(\pi, (x, y_0, \delta)) \geq K_0 \) and \( K_0 = K_a \), otherwise rejected without test if \( R(\pi, (x, y_0, \delta)) \geq K_0 \) and \( K_0 = K_r \).

In practice, engineering experience or manufacturing knowledge are helpful to determine the parameters of prior distribution. The values of \( y_0 \) can be determined as follows: It can be shown that \( E(\theta | \alpha_1, \ldots, \alpha_k) = \alpha_j / \alpha, j = 1, \ldots, k \). Let \( p_j \) be the expected proportion that we expect that \( \lfloor p_j (N-n) \rfloor \) items will fail in the \( (\tau_{j-1}, \tau_j] \), where \( \lfloor x \rfloor \) is the largest integer less than or equal to \( x \). So we can specify the values of \( p_j \) in advance based on our expectation for \( p_j = E(\theta | \alpha_1, \ldots, \alpha_k) \). Assume that \( y_0 \) is a vector of tolerance for satisfied quality requirements, practitioners can specify \( y_{ij} = \lfloor p_j (N-n) \rfloor \) which is the number of failed items in the intervals \( (\tau_{j-1}, \tau_j] \).
i = 1,⋯,k, where 0 < p_i < ⋯ < p_k < 1. An algorithm is provided to determine the optimal BSP as the following steps.

Step 1: Specify the values of τ, α_i, i = 1,⋯,k + 1 for given lot size N and cost components c_i, c_i, and γ.

Step 2: Set n = 1 and the initial loss cost R_0 = ∞.

Step 3: Compute the average loss of accepting and rejecting the lot without sample plan, K_0 and K_γ, respectively. Let K_0 = min{K_0, K_γ}, p_j = α_j / α, and y_j = p_j(N - n), j = 1,⋯,k.

Step 4: Compute the Bayes risk R(π, y_0, δ) by equation (2), and let the nth step risk R_n = R(π, n, y_0, δ).

Step 5: If R_n < R_0, let R_0 = R_n, n = n + 1, and go to Step 4. Otherwise go to Step 6.

Step 6: Let R^* = R_0, n^* = n - 1. If R^* < K_0, the optimal BSP is (n^*, y_0, δ_π(x | n^*, y_0)) and the corresponding Bayes risk is R^*, the optimal sampling number of the lot is n^*. Otherwise, the optimal decision is to accept the lot without test if K_0 = K_δ; and to reject the lot otherwise. Then stop the searching procedure.

3. An Example and Numerical Study

3.1. An Example

Mann and Fertig (1973) gave failure times of airplane components taken during the early development of the F-100 fighter aircraft. The data represent failure times in hours of components aboard the aircraft. Censored samples of ten failures from a total of 13 components subjected to some special test conditions are recorded. The test was censored at the tenth failure time t_10; 13 = 3.00. Engineering considerations indicated that the data should come from a Weibull distribution with pdf:

\[ f(x; \lambda_1, \lambda_2) = \lambda_2 / \lambda_1 \left( t / \lambda_1 \right)^{\lambda_2 - 1} \exp \left[ - \left( t / \lambda_1 \right)^{\lambda_2} \right], t > 0, \]

where \( \lambda_1 \) and \( \lambda_2 \) are scale parameter and shape parameter, respectively. It follows that \( X = \log T \) has an extreme value distribution with the pdf

\[ f(x; u, b) = 1/b \cdot \exp[\{x - u \}/b] \exp \{ - \exp(\{x - u \}/b) \}, -\infty < x < +\infty, \]

where \( u( -\infty < u < +\infty) \) and \( b > 0 \) are parameters. Based on this data, the maximum likelihood estimates of \( u \) and \( b \) are suggested by Mann and Fertig (1973) as \( \hat{u} = 0.821 \) and \( \hat{b} = 0.706 \). The maximum likelihood estimates of \( \lambda_1 \) and \( \lambda_2 \) can then be found as \( \hat{\lambda}_1 = e^{0.821} = 2.273 \) and \( \hat{\lambda}_2 = 1/0.706 = 1.416 \).

Assume that such components are manufactured and packaged in lot each of size 100. We reproduce 100 lifetimes of such components from the Weibull distribution with scale parameter 2.273 and shape parameter 1.416 for the interval censoring test. A lot is sentenced as a good lot if it contains at most 14.3% (\( p_1 = 0.143 \)) items having lifetimes shorter than 1, 28.6% (\( p_2 = 0.286 \)) items having lifetimes within the interval (1,2], and at least 57.1% (1 - \( p_1 - p_2 = 0.571 \)) items having lifetimes longer than 2. The inspected times are specified at \( \tau_1 = 1 \) and \( \tau_2 = 2 \). All survival items at the termination time are censored. Assume that the cost per sampled item is \( c_s = 0.5 \), and the cost per unit time used for testing is \( c_t = 0.7 \). Using the conditions \( p_j = \alpha_j / (\alpha_1 + \alpha_2 + \alpha_3) \), \( j = 1,2,3 \), we have \( \alpha_j = 0.5 \alpha_{j+1}, j = 1,2 \). Taking \( \alpha_1 = 0.5 \), we have \( \alpha_2 = 1 \) and \( \alpha_3 = 2 \). The value of prior parameters may not be specified properly in the initial state of life-testing. But they can be adjusted and brought back to proper values when lifetime data are cumulated. The loss per un-sampled item incurred in a reject lot is considered as \( \gamma = 0.5 \). Table 1 shows that a random sample of 9 components is needed to be
drawn from the lot and subjected to a life test.

Because practitioners only checked failure numbers at the time \( \tau_1 \) and \( \tau_2 \), respectively, only the failure numbers \( x_1 = 1 \) and \( x_2 = 7 \) are observed. Based on the information of \( x = (x_1, x_2) = (1, 7) \) and \( y_0 = (9, 12) \), we have \( H(x) = 1.0 \) which is larger than \( \gamma = 0.5 \). We conclude to reject the lot.

Table 1: Optimal BSPs for \( \alpha_1 = 0.5, \alpha_2 = 1, \alpha_3 = 2, c_s = 0.5, c_r = 0.7, \tau = 1 \) and \( \gamma = 0.5 \) when \( k = 2 \).

<table>
<thead>
<tr>
<th>( N )</th>
<th>( n^* )</th>
<th>( y_0 )</th>
<th>( R^* )</th>
<th>( K_a )</th>
<th>( K_r )</th>
<th>( K_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>8</td>
<td>(5,11)</td>
<td>24.279</td>
<td>36.482</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>100</td>
<td>9</td>
<td>(9,12)</td>
<td>45.984</td>
<td>72.322</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>200</td>
<td>11</td>
<td>(26,53)</td>
<td>88.309</td>
<td>144.055</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>400</td>
<td>18</td>
<td>(54,108)</td>
<td>169.858</td>
<td>288.037</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>600</td>
<td>26</td>
<td>(81,163)</td>
<td>248.903</td>
<td>431.048</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>800</td>
<td>37</td>
<td>(108,217)</td>
<td>325.779</td>
<td>574.548</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>1000</td>
<td>41</td>
<td>(136,273)</td>
<td>404.327</td>
<td>718.044</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>

3.2. Numerical Study

A numerical study is conducted with three equally-spaced intervals \( k = 2 \) to illustrate the proposed method, that is, we use the conditions of \( \tau_j - \tau_{j-1} = \tau \) to develop BSPs. The scenario is to determine the BSPs for lot sizes under the consideration of administration convenience. Using the algorithm in Section 2, some optimal BSPs are tabulated for \( N = 50,100,200,400,600,800,1000 \), given \( \alpha_1 = 0.5, \alpha_2 = 1, \alpha_3 = 2, c_s = 0.5, c_r = 0.7, \tau = 1 \) and \( \gamma = 0.5 \). We expect a satisfied lot of item should contain at most 14.3% items having lifetimes shorter than \( \tau \), 28.6% items having lifetimes within the interval \( (\tau, 2\tau] \), and at least 57.1% items having lifetimes longer than \( 2\tau \). Without lost of generality, we take \( \tau = 1 \) as the unit test time. Programs were written in R language. The R code can be obtained from the author upon request. All the numerical results are given in Table 1, which indicates that the Bayes risk and sample size increase as the lot size increases.

For example, if items are manufactured in lots each of size \( N = 800 \). An interval censoring test with three equally-spaced intervals is prepared for the inspection of acceptance sampling. Based on the manufacturing condition, Table 1 indicates that the BSPs suggest a random sample of 37 items needed to be drawn from the lot and put on a life test. Assume that the unit test time is 250. Numbers of failed items in \( (\tau_1, \tau_2) \) and \( (\tau_2, \tau_3) \) should be recorded as \( x_1 \) and \( x_2 \), respectively, and the survival items at the termination time, the 500th hour, is counted as \( x_3 \). The value of \( H(x) \) of (4) is computed. The lot of items is accepted if \( H(x) - \gamma < 0 \), and rejected otherwise. For instance, based on the information of \( x = (x_1, x_2) = (2,8) \) and \( y_0 = (109,218) \), we have \( H(x) = 0.477 \) which is smaller than \( \gamma = 0.5 \). We conclude to accept the lot. However, if the information of \( x = (x_1, x_2) = (8,2) \), \( y_0 = (109,218) \) and \( H(x) = 0.861 \) which is larger than \( \gamma = 0.5 \), we conclude to reject the lot.

4. Conclusions

In this article, a life testing procedure with interval censoring is established based on the Bayesian approach. The interval probabilities are assumed to be random and vary from lot to lot according to a Dirichlet distribution. A suitable loss function is taken including the cost per sampled item, the cost per unit time used and the decision loss to determine the Bayes risk. An algorithm of finding optimal Bayesian sampling plans.
is presented. An example is used and a numerical study is conducted to illustrate the proposed method. Moreover, a sensitivity analysis study is conducted to evaluate the influences of prior parameters and sampling costs on the proposed sampling plan.

The study shows that the Bayes risk and sample size increase when the lot size increases. The sensitivity analysis also indicates that the sample size changes irregularly as the prior parameters change. The change in Bayes risk is not significant when the prior parameters change. The Bayes risk increases and sample size decreases as the cost per sampled item increases. The proposed sampling procedure only depends on the prior information and the cost components selection, but frees of the life distribution of items. So the proposed method can be used for various life distributions. Extend the proposed method to empirical Bayesian approach would be investigated in future.

Acknowledgements
This paper was supported by National Natural Science Foundation of China (Grant No. 71101118).

References